ON THE FUNCTIONAL CLT FOR PARTIAL SUMS OF TRUNCATED BOUNDED FROM BELOW RANDOM VARIABLES

VLADIMIR POZDNYAKOV

Abstract. Let $\{X, X_i\}_{i\geq 1}$ be i.i.d. bounded from below continuous random variables, $\mathbf{E}|X| < \infty$, $\mathbf{E}X^2 = \infty$, and $\{b_n\}_{n\geq 1}$ be a sequence of increasing positive numbers. When X belongs to the Feller class and b_n is such that $n\mathbf{P}(X > b_n) \to \infty$ and $\mathbf{E}(X\mathrm{I}_{X>b_n}) \big/ \mathbf{E}(X^2\mathrm{I}_{X\leq b_n}) = o((\ln n)^{-1})$, a functional CLT for the truncated sums $S_n = \sum_{i=1}^{n} X_i\mathrm{I}_{X_i\leq b_n}$ is proved.
Keywords. Functional CLT, Truncated sums, trimmed sums, Martingale.

1. Introduction

Assume that we observe sequence of i.i.d. positive random variables with a finite mean but an infinite variance. By the Kolmogorov's strong law of large numbers the sample mean converges to the true mean with probability 1. But the classic Central Limit Theorem (CLT) is not valid any longer. What we can do is to exclude the observations that exceed some given deterministic level. If the truncating level increases to infinity as sample size is getting larger then the sample mean of truncated observations is still a.s. consistent estimator of the true mean. But in this case one also can show that under quite general conditions on the distribution and truncating level we have an analog of the classic CLT. In fact, the proof of this result is relatively simple. It only requires an application of some standard results from the limit theory of sequences of series of independent random variables, because for any *fixed* sample size we have just a sum of i.i.d. bounded random variables.

However, the functional CLT for truncated sums presents a more challenging problem due to the fact that the sequence of truncated sums is not a well-behaved process. Perhaps this is one of the reasons why there are not many functional limit theorems proven for truncated (or trimmed) sums. The list of related results is Ould-Rois (1991), Kasahara (1993), Egorov and Pozdnyakov (1997) and Pozdnyakov (2003). But exactly this kind of limit theorem one needs to use, for example, in the sequential analysis (Sen (1981)) which means that the presented result could be also useful in applications.

The method that we use here is as follows. For each n we present the truncated sum as a sum of two terms. The first terms are small in some appropriate sense, while the sequence of the second terms forms a martingale. The functional CLT for the martingale is proved using a scheme that was proposed in Pozdnyakov (2003) for the case of truncated sums of symmetric random variables.

Date: April 4, 2006.
Address: Department of Statistics, University of Connecticut, 215 Glenbrook Road, U-4120, Storrs, CT 06269-4120, *email:* boba@stat.uconn.edu.

Another interesting aspect of this result is that the considered distributions are not symmetric. If truncating (or trimming) is based on the absolute values of random variables then the symmetry of the distribution plays an important role. Many results are obtained under the symmetry assumption. See, for example, Pruitt (1988), Griffin and Pruitt (1987), Hahn and Kuelbs (1989), Haeusler and Mason (1990), Hahn et al (1991), Cuzick et al (1995) and Griffin and Qazi (2002). However, if the distribution is not symmetric many problems remain open. We believe that the martingale approach employed here could be useful in solving those open problems.

2. Main Result

Let $\{X, X_i\}_{i\geq 1}$ be i.i.d random variables with an continuous distribution and infinite second moment. Assume that

$$\mathbf{E}X = 0, \tag{1}$$

and

$$X > -M \text{ a.s.} \tag{2}$$

where $M > 0$. Without loss of generality we can shift random variables by their expected values. This will simplify the notation.

Assume also that the random variable X belongs to the Feller class, i.e.

$$\limsup_{t\to\infty} \frac{t^2\mathbf{P}(X > t)}{\mathbf{E}\big(X^2 \mathrm{I}_{X\leq t}\big)} < \infty. \tag{3}$$

Let $\{b_n\}_{n\geq 1}$ be a sequence of positive numbers such that $b_n \uparrow \infty$. The *truncated sum* S_n we will consider are define by

$$S_n = \sum_{i=1}^{n} X_i \mathrm{I}_{X_i\leq b_n}. \tag{4}$$

The main result is a functional CLT for S_n. Let d_n be the smallest (negative) number that

$$\mathbf{E}\big(X \mathrm{I}_{d_n\leq X\leq b_n}\big) = 0, \tag{5}$$

$E_n = \{x : d_n \leq x \leq b_n\}$, and $A_n = n\mathbf{E}(X^2\mathrm{I}_{X\in E_n})$. Since the distribution is continuous the number always exists and $d_n \downarrow$ as $n \to \infty$. Let $S_n(t)$ be a random element of $\mathcal{C}[0, 1]$ obtained by linear interpolation between the points

$$\Big(0, 0\Big), \Big(\frac{A_1}{A_n}, \frac{S_1 - \mathbf{E}S_1}{\sqrt{A_n}}\Big), ..., \Big(1, \frac{S_n - \mathbf{E}S_n}{\sqrt{A_n}}\Big).$$

Theorem 1. *If the random variable X satisfies conditions (1), (2) and (3), the average number of the excluded variables*

$$n\mathbf{P}\big(X > b_n\big) \to \infty, \tag{6}$$

b_n grows fast enough to guarantee that the ratio

$$\frac{\mathbf{E}(X\mathrm{I}_{X>b_n})}{\mathbf{E}(X^2\mathrm{I}_{X\leq b_n})} = o((\ln n)^{-1}), \tag{7}$$

and

$$\frac{A_n}{A_{n+1}} \to 1, \tag{8}$$

then

$$S_n(t) \xrightarrow{d} W \tag{9}$$

in the sense $\mathcal{C}[0,1]$ *with uniform metric* ρ *where* W *is standard Brownian motion on* $[0,1]$.

3. Proof of the Main Result

First let us note that for a fixed n the truncated sums S_n is a sum of independent identically distributed bounded random variables. Thus, to establish a central limit theorem for S_n we just need to check standard conditions say as it is given by Petrov (1995, p. 113). Let

$$B_n = \sum_{j=1}^{n} \mathbf{E}\big(X_j^2 \mathrm{I}_{X_j \le b_n}\big) = n\mathbf{E}\big(X^2 \mathrm{I}_{X \le b_n}\big). \tag{10}$$

The key observation is that conditions (3) and (6) imply that

$$b_n = o(B_n).$$

Keeping this in mind one can prove the following result.

Theorem 2. *Assume the random variable* X *and sequence* b_n *satisfy conditions (2), (3), and (6).*

Then

$$\frac{S_n - n\mathbf{E}\big(X \mathrm{I}_{X \le b_n}\big)}{\sqrt{B_n}} \xrightarrow{d} \mathcal{N}(0,1).$$

This gives us a base to believe that the functional CLT is also valid under similar conditions. However, in the case of functional limit theorem we have to think about sequence $\{S_n\}_{n\ge1}$ as a *process*. Due to the truncation $\{S_n\}_{n\ge1}$ is not a process with independent increments. It is also not monotone. As a consequence, many needed tools (for instance, Kolmogorov's inequality) are not available. The main trick is to present process $\{S_n\}$ as a sum of two well-behaved processes.

More specifically, denote $R_n = \sum_{j=1}^n X_j \mathrm{I}_{X_j < d_n}$, and $M_n = \sum_{j=1}^n X_j \mathrm{I}_{X_j \in E_n}$, so $S_n = R_n + M_n$. First we show that the M_n is a martingale with respect to a certain σ-field. Using this fact we establish then a weak invariance principle for M_n. Secondly, we show that R_n, in a sense, can be ignored.

Lemma 1. *Let* $\mathcal{F}_n = \sigma\big(X_1 \mathrm{I}_{X_1 \in E_n}, ..., X_n \mathrm{I}_{X_n \in E_n}\big)$. *The sequence* $\{M_n, \mathcal{F}_n\}_{n>0}$ *with* $M_0 = 0$ *and* $\mathcal{F}_0 = \{\Omega, \emptyset\}$ *is a martingale, and its predictable quadratic variation* $\langle M\rangle_n = \sum_{i=1}^n \mathbf{E}\big(\big(M_i - M_{i-1}\big)^2 | \mathcal{F}_{i-1}\big)$ *is given by*

$$\langle M\rangle_n = \sum_{i=1}^{n} \mathbf{E}\big(X^2 \mathrm{I}_{X \in E_i}\big) + \sum_{i=2}^{n} \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X \in E_i \setminus E_{i-1}}\big)}{\mathbf{P}\big(X \notin E_{i-1}\big)} \sum_{j=1}^{i-1} \mathrm{I}_{X_j \notin E_{i-1}}. \tag{11}$$

Proof. It is clear that $\{\mathcal{F}_n\}$ is a filtration, and M_n is $\mathcal{F}_n$-measurable. Now note that

$$M_{n+1} - M_n = X_{n+1} \mathrm{I}_{X_{n+1} \in E_{n+1}} + \sum_{j=1}^{n} X_j \mathrm{I}_{X_j \in E_{n+1} \setminus E_n}.$$

Using Lemmas 1, 2 and 3 of Pozdnyakov (2003) one can show then that

$$\mathbf{E}\big(X_{n+1} \mathrm{I}_{X_{n+1} \in E_{n+1}} \big| \mathcal{F}_n\big) = \mathbf{E}\big(X_{n+1} \mathrm{I}_{X_{n+1} \in E_{n+1}}\big) = 0,$$

and for $j < n$

$$\mathbf{E}\big(X_i \mathrm{I}_{X_i \in E_{n+1}\setminus E_n} \big| \mathcal{F}_n\big) = \frac{\mathbf{E}\big(X_i \mathrm{I}_{X_i \in E_{n+1}\setminus E_n}\big)}{\mathbf{P}\big(X_i \notin E_n\big)} \mathrm{I}_{X_i \notin E_n} = 0.$$

This proves that $\{M_n, \mathcal{F}_n\}_{n>0}$ is a martingale. As to the predictable quadratic variation, it is not difficult to find that

$$\mathbf{E}\big(\big(M_i - M_{i-1}\big)^2 | \mathcal{F}_{i-1}\big) = \mathbf{E}\big(X_i^2 \mathrm{I}_{X_i \in E_i}\big) + \sum_{j=1}^{i-1} \frac{\mathbf{E}\big(X_j^2 \mathrm{I}_{X_j \in E_i \setminus E_{i-1}}\big)}{\mathbf{P}\big(X_j \notin E_{i-1}\big)} \mathrm{I}_{X_j \notin E_{i-1}}$$

$$= \mathbf{E}\big(X^2 \mathrm{I}_{X \in E_i}\big) + \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X \in E_i \setminus E_{i-1}}\big)}{\mathbf{P}\big(X \notin E_{i-1}\big)} \sum_{j=1}^{i-1} \mathrm{I}_{X_j \notin E_{i-1}}.$$

This finishes the proof of the lemma. □

Now we are ready to formulate the functional CLT for M_n The variance of the martingale M_n is A_n. It is clear that

$$A_n = \mathbf{Var}(M_n) = \mathbf{E}\langle M \rangle_n = \sum_{j=1}^{n} \mathbf{E}\big(X_j^2 \mathrm{I}_{X_j \in E_n}\big) = n\mathbf{E}(X^2 \mathrm{I}_{X \in E_n}) \sim B_n.$$

Let $M_n(t)$ be a random element of $\mathcal{C}[0,1]$ obtained by linear interpolation between the points $(0,0)$, $(A_1/A_n, M_1/\sqrt{A_n})$,..., $(1, M_n/\sqrt{A_n})$. More specifically,

$$M_n(t) = \frac{1}{\sqrt{A_n}} \left[M_i + \frac{tA_n - A_i}{(A_{i+1} - A_i)} (M_{i+1} - M_i) \right], \quad \frac{A_i}{A_n} \le t < \frac{A_{i+1}}{A_n}.$$

Brown (1971) showed that in order to establish the functional CLT for truncated sum S_n we need to verify the Lindeberg condition

$$\text{for all } \epsilon > 0, \quad \frac{1}{A_n} \sum_{i=1}^{n} \mathbf{E}\big((M_i - M_{i-1})^2 \mathrm{I}_{|M_i - M_{i-1}| > \epsilon\sqrt{A_n}}\big) \to 0, \quad \text{as } n \to \infty \tag{12}$$

and the weak law of large numbers for the predictable quadratic variation $\langle M \rangle_n$

$$\frac{\langle M \rangle_n}{A_n} \xrightarrow{\mathbf{P}} 1. \tag{13}$$

If conditions (12) and (13) hold, then

$$M_n(t) \xrightarrow{d} W. \tag{14}$$

The proof of the next result employs an approach similar to the one used in Pozdnyakov (2003) for symmetric random variables. The random variables X_n are not symmetric here, but our special choice of truncating levels b_n and d_n allows us to use the same scheme with some obvious adjustments, so we omit the proof.

Proposition 1. *If the random variable X satisfies conditions (1), (2), (3), (6) and (8) then conditions (12) and (13) hold, and, as a consequence, $M_n(t) \xrightarrow{d} W$ in the sense $(\mathcal{C}[0,1], \rho)$.*

Note that

$$S_n - \mathbf{E}S_n = R_n - \mathbf{E}R_n + M_n.$$

Thus, if we show that $|R_n - \mathbf{E}R_n|$ are small, then we can substitute M_n in Proposition 1 by S_n to get Theorem 1. More specifically, it would be sufficient if we prove the following lemma.

Lemma 2. *Under the conditions of Theorem 1*

$$\max_{i\le n}\frac{|R_i-\mathbf{E}R_i|}{\sqrt{A_n}}\xrightarrow{P}0. \tag{15}$$

Proof. First let us show that the variance of R_n is small, indeed. Just note that

$$\begin{aligned}\mathbf{Var}(R_n)&=n\mathbf{Var}(X\mathrm{I}_{X<d_n})\\&\le n\mathbf{E}(X^2\mathrm{I}_{X<d_n})\\&\le nM\mathbf{E}(|X|\mathrm{I}_{X<d_n})\\&=nM\mathbf{E}(X\mathrm{I}_{X>b_n}).\end{aligned}$$

Therefore,

$$\mathbf{Var}\left(\frac{R_n}{\sqrt{B_n}}\right)=O\left(\frac{\mathbf{E}(X\mathrm{I}_{X>b_n})}{\mathbf{E}(X^2\mathrm{I}_{X\le b_n})}\right), \tag{16}$$

which goes to 0 as b_n gets larger. This is enough to show that

$$\frac{|R_n-\mathbf{E}R_n|}{\sqrt{B_n}}\xrightarrow{P}0.$$

But we need a bit more. To prove (15) we will show first that under the conditions of Theorem 1 we have *almost sure* convergence for R_n, i.e with probability one

$$\frac{|R_n-\mathbf{E}R_n|}{\sqrt{B_n}}\longrightarrow\quad 0. \tag{17}$$

This could be proved via the straightforward application of the Bernstein inequalities (see, for example, Petrov (1995, p. 57). Since $|X\mathrm{I}_{X<d_n}-\mathbf{E}X\mathrm{I}_{X<d_n}|$ is a.s. bounded by $2M$ we get that for any $\epsilon>0$ we have the following estimate:

$$\mathbf{P}(|R_n-\mathbf{E}R_n|>\epsilon\sqrt{B_n})\le 2\max\{e^{-\epsilon^2B_n/(4\mathbf{Var}(R_n))},e^{-\epsilon\sqrt{B_n}/(8M)}\}.$$

If sequence $\{b_n\}$ satisfies the condition (7), equation (16) tells us that

$$e^{-\epsilon^2B_n/(4\mathbf{Var}(R_n))}=o\left(\frac{1}{n^2}\right).$$

On the other hand,

$$\frac{n}{B_n}\to 0.$$

Therefore, we get that

$$\sum_{n=1}^{\infty}\mathbf{P}(|R_n-\mathbf{E}R_n|>\epsilon\sqrt{B_n})<\infty.$$

The Borel-Cantelli lemma gives us a.s. convergence. It easy to see that (17) implies that

$$\max_{i\le n}\frac{|R_i-\mathbf{E}R_i|}{\sqrt{B_n}}\longrightarrow 0 \text{ a.s.} \tag{18}$$

Indeed, for each $\epsilon>0$ and almost all ω one can find $n_1=n_1(\epsilon,\omega)$ such that for all $n>n_1$

$$\frac{|R_n-\mathbf{E}R_n|}{\sqrt{B_n}}<\epsilon.$$

Since $B_n \uparrow \infty$ there exists $n_2 = n_2(\epsilon, \omega) > n_1$ that for all $n > n_2$

$$\frac{\max_{i<n_1} |R_i - \mathbf{E}R_i|}{\sqrt{B_n}} < \epsilon,$$

and this proves (18). Since $B_n \sim A_n$ we finally get (15). □

4. Concluding Remarks

As we can see moment condition (1) is not needed for the ordinary CLT, so it is possible that this condition could be also dropped in the case of the functional CLT. But at the moment we do not now how to do it. Nevertheless, the class of distributions that fit the description given in Theorem 1 is quite large. For example, positive random variables that belongs to the domain of attraction of a stable distribution with $1 < \alpha < 2$ are in the considered class. These distributions are often used to model situations when there are outliers in data. So, the functional CLT presented here could be potentially used to design sequential procedures for continuous monitoring of such data.

Conditions (6) and (7) essentially determine the rate of growth of the sequence b_n. While (6) is required even for the ordinary CLT, condition (7) is more technical. We need it to produce an estimate related to R_n – not a nice process to work with. But they are not very restrictive in practical applications. Again, let us consider the case when X belongs to the domain of attraction of a stable law with $1 < \alpha < 2$, i.e. there exists a slowly varying function $L(x)$ such that

$$\mathbf{E}(X^2, X < x) = x^{2-\alpha} L(x),$$

$$\mathbf{P}(X > x) \sim \frac{2-\alpha}{\alpha} x^{-\alpha} L(x),$$

and

$$\mathbf{E}(X, X > x) \sim \frac{2-\alpha}{\alpha - 1} x^{1-\alpha} L(x),$$

as $x \to +\infty$. Let $b_n = cn^\gamma$. What γ do we need to satisfy conditions of Theorem 1? Condition (6) (the one that we need even for the ordinary CLT) produces $\gamma < 1/\alpha$. Condition (7) requires $\gamma > 0$ which is not really a restriction for the particular choice of b_n. We need this anyway to get an increasing sequence. If $b_n = c(\ln n)^\gamma$, then we need nothing to get (6), and we need $\gamma > 1$ to have (7).

Finally, let us note that condition (8) looks technical, but, in fact, it is impossible to omit this one. Indeed, since the variance A_n goes to infinity faster than n, by choosing b_n that changes by occasional jumps we can easily make $S_n(t)$ to be very different from the Brownian motion.

References

[1] Brown, B.M. (1971), Martingale central limit theorems, *Ann. Math. Stat.*, **42**, 59-66.

[2] Cuzick, J., Giné, E., and Zinn, J. (1995), Laws of large numbers for quadratic forms, maxima of products and truncated sums of i.i.d. random variables, *Ann. Probab.*, **23**, 292-333.

[3] Egorov, V. and Pozdnyakov, V. (1997), Functional law of the iterated logarithm for trimmed sums, (Russian) *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **244**, Veroyatn. i Stat. 2, 126-142; translation in J. Math. Sci. (New York) **99** (2000) 1094-1104.

[4] Haeusler, E. and Mason, D. (1990), A law of the iterated logarithm for modulus trimming. *Colloq. Math. Soc. Janos Bolyai*, **57**, 241-254.

[5] Hahn, M. and Kuelbs, J. (1989), Universal asymptotic normality for conditionally trimmed sums, *Stat. & Prob. Letters*, **7**, 9-15.

[6] Hahn, M., Kuelbs, J., and Weiner, D. (1991), Asymptotic behavior of partial sums: a more robust approach via trimming and self-normalization, *Sum, trimmed sums and extremes, Progr. Probab.*, **23**, 1-53.

[7] Griffin, P. and Qazi, F. (2002), Limit laws of modulus trimmed sums, *Ann. Probab.*, **30**, 1466-1485.

[8] Griffin, P. and Pruitt, W. (1987), The central limit theorem for trimmed sums, *Math. Proc. Cambridge Philos. Soc.*, **102**, 329-349.

[9] Kasahara, Y. (1993), A functional limit theorem for trimmed sums, *Stochastic Process. Appl.*, **47**, 315-322.

[10] Ould-Rois, H. (1991), Invariance principles and self-normalizations for sums trimmed according to choice of influence function, *Sum, trimmed sums and extremes, Progr. Probab.*, **23**, 55-80.

[11] Petrov, V.V. (1995), *Limit theorems of probability theory. Sequences of independent random variables* (Oxford University Press, New York)

[12] Pozdnyakov, V. (2003), A Note on functional CLT for truncated sums, *Statistics & Probability Letters*, **61**, 277-286.

[13] Pruitt, W. (1988), Sums of independent random variables with the extreme terms excluded. *Probability and statistics. Essays in honor of Franklin A. Graybill (J.N. Srivastava, Ed.)*, 201-216.

[14] Sen, P. K. (1981), *Sequential Nonparametrics: Invariance Principles and Statistical Inference*, Jonh Wiley & Sons, New York.

Appendix A: Proof of the Theorem 2

Let

$$X_{nj} = \frac{X_j \mathrm{I}_{X_j \le b_n} - \mathbf{E}\big(X \mathrm{I}_{X \le b_n}\big)}{\sqrt{B_n}}.$$

According to Petrov (1995, p. 113) to establish the CLT for this sequence of series of random variables we need to check that for every fixed $\epsilon > 0$ the following conditions hold

$$\sum_{j=1}^{n} \mathbf{P}\big(|X_{nj}| \ge \epsilon\big) \to 0,$$

$$\sum_{j=1}^{n} \mathbf{Var}\big(X_{nj} \mathrm{I}_{|X_{nj}|<\epsilon}\big) \to 1,$$

$$\sum_{j=1}^{n} \mathbf{E}\big(X_{nj} \mathrm{I}_{|X_{nj}|<\epsilon}\big) \to 0.$$

First note that conditions (3) and (6) imply that

$$b_n = o(B_n).$$

Therefore, for all sufficiently large n we have that $|X_{jn}| < \epsilon$ with probability 1 and, as consequence,

$$\sum_{j=1}^{n} \mathbf{P}\big(|X_{nj}| \ge \epsilon\big) = n\mathbf{P}\big(|X \mathrm{I}_{X \le b_n} - \mathbf{E}\big(X \mathrm{I}_{X \le b_n}\big)| > \epsilon B_n\big) = 0,$$

$$\sum_{j=1}^{n} \mathbf{Var}\big(X_{nj} \mathrm{I}_{|X_{nj}|<\epsilon}\big) = \sum_{j=1}^{n} \mathbf{Var}\big(X_{nj}\big) = \frac{\mathbf{Var}(S_n)}{B_n},$$

and

$$\sum_{j=1}^{n} \mathbf{E}\big(X_{nj} \mathrm{I}_{|X_{nj}|<\epsilon}\big) = \sum_{j=1}^{n} \mathbf{E}\big(X_{nj}\big) = 0.$$

But since $\mathbf{E}X^2 = \infty$ we have that

$$\frac{\mathbf{Var}(S_n)}{B_n} \to 1.$$

Appendix B: Proof of the Proposition 1

Let us first to show that

$$\mathbf{E}\left(\frac{\langle M \rangle_n - A_n}{A_n}\right)^2 \to 0, \tag{19}$$

then, as consequence, we get trivially (13). If we define α_i by

$$\alpha_i = \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X \in E_{i+1} \setminus E_i}\big)}{\mathbf{P}(X \notin E_i)},$$

then the predictable quadratic variation is given by

$$\begin{aligned}\langle M\rangle_n = \quad & \mathbf{E}\big(X_1^2 \mathrm{I}_{X_1\in E_1}\big) && +\mathbf{E}\big(X_2^2 \mathrm{I}_{X_2\in E_2}\big) && +... && +\mathbf{E}\big(X_n^2 \mathrm{I}_{X_n\in E_n}\big)\\ & +\alpha_1 \mathrm{I}_{X_1\notin E_1} \\ & +\alpha_2 \mathrm{I}_{X_1\notin E_2} && +\alpha_2 \mathrm{I}_{X_2\notin E_2}\\ & +... \\ & +\alpha_{n-1} \mathrm{I}_{X_1\notin E_{n-1}} && +\alpha_{n-1} \mathrm{I}_{X_2\notin E_{n-1}} && +... && +0.\end{aligned}$$

Therefore, the quadratic variation can be viewed as a sum of independent random variables. Specifically,

$$\langle M\rangle_n = \sum_{i=1}^{n} \mathbf{E}\big(X_i^2 \mathrm{I}_{X_i\in E_i}\big) + \sum_{i=1}^{n-1} Y_{in},$$

where

$$\begin{aligned}Y_{in} =& \alpha_i \mathrm{I}_{X_i\notin E_i} + ... + \alpha_{n-1}\mathrm{I}_{X_i\notin E_{n-1}}\\ =& \alpha_i \mathrm{I}_{X_i\in E_{i+1}\setminus E_i} + (\alpha_i+\alpha_{i+1})\mathrm{I}_{X_i\in E_{i+2}\setminus E_{i+1}} + ... + (\alpha_i + ... + \alpha_{n-1})\mathrm{I}_{X_i\notin E_{n-1}}.\end{aligned}$$

Now note that

$$\mathbf{E}Y_{in} = \mathbf{E}(X^2 \mathrm{I}_{X\in E_n\setminus E_i})$$

and

$$Y_{in} \le \alpha_i + ... + \alpha_{n-1} \quad a.s.$$

These two observations lead us to the following inequality:

$$\mathbf{Var}(Y_{in}) \le \mathbf{E}[Y_{in}]^2 \le (\alpha_i + ... + \alpha_{n-1})\mathbf{E}Y_{in}.$$

Since

$$\begin{aligned}\alpha_i + ... + \alpha_{n-1} =& \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X\in E_{i+1}\setminus E_i}\big)}{\mathbf{P}(X\notin E_i)} + ... + \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X\in E_n\setminus E_{n-1}}\big)}{\mathbf{P}(X\notin E_{n-1})}\\ \le& \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X\in E_{i+1}\setminus E_i}\big)}{\mathbf{P}(X\notin E_{n-1})} + ... + \frac{\mathbf{E}\big(X^2 \mathrm{I}_{X\in E_n\setminus E_{n-1}}\big)}{\mathbf{P}(X\notin E_{n-1})}\\ =& \frac{\mathbf{E}(X^2 \mathrm{I}_{X\in E_n\setminus E_i})}{\mathbf{P}(X\notin E_{n-1})}\\ \le& \frac{\mathbf{E}(X^2 \mathrm{I}_{X\in E_n\setminus E_i})}{\mathbf{P}(X> b_n)}\\ =& \frac{\mathbf{E}Y_{in}}{\mathbf{P}(X>b_n)},\end{aligned}$$

we get that

$$\mathbf{Var}(Y_{in}) \le \frac{[\mathbf{E}Y_{in}]^2}{\mathbf{P}(X>b_n)} \le \frac{A_n^2}{n^2\mathbf{P}(X>b_n)}.$$

Thus, finally we get

$$\mathbf{E}\left(\frac{\langle M\rangle_n - A_n}{A_n}\right)^2 = \frac{\mathbf{Var}(\sum_{i=1}^{n-1} Y_{in})}{A_n^2} \le \frac{1}{A_n^2}\frac{nA_n^2}{n^2\mathbf{P}(X>b_n)} = \frac{1}{n\mathbf{P}(X>b_n)} \to 0.$$

Now let us show that the martingale M_n satisfies Lindenberg condition (12). By the Cauchy inequality we have

$$\begin{aligned}&\mathbf{E}\big((M_i - M_{i-1})^2 \mathrm{I}_{|M_i-M_{i-1}|>\epsilon\sqrt{A_n}}\big)\\ &\qquad \le \big(\mathbf{E}(M_i-M_{i-1})^4\big)^{1/2}\big(\mathbf{P}\big(|M_i-M_{i-1}|>\epsilon\sqrt{A_n}\big)\big)^{1/2}.\end{aligned}$$

First note that by the Chebyshev inequality we have

$$\mathbf{P}\big(|M_i - M_{i-1}| > \epsilon\sqrt{A_n}\big) \le \frac{A_i - A_{i-1}}{\epsilon^2 A_n}.$$

Now let us estimate $\mathbf{E}(M_i - M_{i-1})^4$. To do this let us first present the martingale difference $M_i - M_{i-1}$ as a sum of independent random variables. Specifically, the martingale difference can be viewed in the following way:

$$M_i - M_{i-1} = \xi_1 + ... + \xi_{i-1} + \xi_i,$$

where $\xi_i = X_i \mathrm{I}_{X_i \in E_i}$ and $\xi_j = X_j \mathrm{I}_{X_j \in E_i \setminus E_{i-1}}$ for $j = 1, 2, ..., i-1$. Note that $\mathbf{E}\xi_j = 0$ for all $j = 1, .., i$, therefore, we get that

$$\begin{aligned}\mathbf{E}(M_i - M_{i-1})^4 =& \sum_{j=1}^{i} \mathbf{E}\xi_j^4 + 6 \sum_{1\le j<k\le i} \mathbf{E}\xi_j^2 \mathbf{E}\xi_k^2 \\ \le& \sum_{j=1}^{i} \mathbf{E}\xi_j^4 + 3\left[\sum_{j=1}^{i} \mathbf{E}\xi_j^2\right]^2\end{aligned}$$

Since for all sufficiently large n

$$\xi_j^2 \le b_i^2 \le b_n^2 \text{ for all } 1 \le j \le i \le n$$

we find that

$$\mathbf{E}(M_i - M_{i-1})^4 \le b_n^2 (A_i - A_{i-1}) + 3(A_i - A_{i-1})^2.$$

Hence, we have that

$$\begin{aligned}\mathbf{E}\big((M_i - M_{i-1})^2 \mathrm{I}_{|M_i - M_{i-1}| > \epsilon\sqrt{A_n}}\big) \le& \frac{\sqrt{b_n^2 + 3(A_i - A_{i-1})}(A_i - A_{i-1})}{\epsilon\sqrt{A_n}} \\ \le& \frac{b_n (A_i - A_{i-1})}{\epsilon\sqrt{A_n}} + \frac{2(A_i - A_{i-1})^{3/2}}{\epsilon\sqrt{A_n}}.\end{aligned}$$

Thus, we get that

$$\frac{1}{A_n} \sum_{i=1}^{n} \mathbf{E}\big((M_i - M_{i-1})^2 \mathrm{I}_{|M_i - M_{i-1}| > \epsilon\sqrt{A_n}}\big) \le$$

$$\begin{aligned}&\le \frac{1}{A_n} \sum_{i=1}^{n} \frac{b_n (A_i - A_{i-1})}{\epsilon\sqrt{A_n}} + \frac{2}{A_n} \sum_{i=1}^{n} \frac{(A_i - A_{i-1})^{3/2}}{\epsilon\sqrt{A_n}} \\ &\le \frac{b_n}{\epsilon A_n^{3/2}} \sum_{i=1}^{n} (A_i - A_{i-1}) + \frac{2}{\epsilon A_n^{3/2}} \sum_{i=1}^{n} (A_i - A_{i-1})^{3/2} \\ &\le \frac{b_n}{\epsilon A_n^{1/2}} + \frac{2}{\epsilon A_n^{3/2}} \sum_{i=1}^{n} (A_i - A_{i-1})^{3/2}.\end{aligned}$$

Because of (3) and (6) we have

$$\frac{b_n}{\sqrt{A_n}} \sim \frac{b_n}{\sqrt{B_n}} = O\left(\frac{1}{\sqrt{n\mathbf{P}(|X| > b_n)}}\right) \to 0$$

as $n \to \infty$, so the first term goes to zero. Finally, it is easy to show that

$$\frac{1}{A_n^{3/2}} \sum_{i=1}^{n} (A_i - A_{i-1})^{3/2} \leq \left(\frac{\max_{i \leq n} A_i - A_{i-1}}{A_n} \right)^{1/2} \to 0,$$

if $A_n/A_{n+1} \to 1$. This finishes the proof of the proposition.